Integration of single-cell transcriptomic and chromatin accessibility on heterogenicity of human peripheral blood mononuclear cells utilizing microwell-based single-cell partitioning technology

Poster
B902

Elham Hatami[1*], Thomas McCarthy[2*], Hye-Won Song[1], Quyen Bao[2], Youngsook Kim[2], Hongduan Huang[2], Rosary Nguyen[2], Punya Narayan[2], Zhiqi Zhang[1], Samuel Shi[1], Ruifang Li[2], Larry Wang[2] and Aruna Ayer[2]

*Authors contributed equally

[1]BD Biosciences, 10975 Torreyana Rd, San Diego, CA 92121, [2]BD Biosciences, 2350 Qume Drive, San Jose, CA, 95131

Abstract

Single-cell RNA sequencing (scRNA-Seq) has significantly advanced our understanding of cell development and heterogeneity in complex systems at cellular resolution. However, its capacity to elucidate cell states and uncover gene regulatory programs remains limited. In contrast, chromatin state profiles serve as a valuable resource for gauging gene expression potential and provide insights regarding transcriptional regulation. When integrated with gene expression data, chromatin accessibility region (CAR) profiles offer a means to define the fundamental gene regulatory logic as the foundation of cell fate.

ATAC-Seq (Assay for Transposase-Accessible Chromatin using Sequencing) is a potent approach for profiling genome-wide CARs. To capture both mRNA gene expression data and gene regulatory information in a single assay, we conducted multiomic scATAC-Seq + scRNA-Seq on PBMCs, utilizing the gentle and robust microwell-based single-cell partitioning technology of the BD Rhapsody™ System. This scATAC-Seq assay showed high sensitivity (many unique fragments of appropriate size per cell) and specificity (most fragments clustered into distinct peaks, with significant enrichment around annotated Transcription Start Sites).

Here we demonstrate the power of this multiomics assay for exploring the regulation of gene expression through chromatin accessibility. Cell types were annotated, and we found enrichment of transcription factor motifs within CARs in each distinct cell type. Gene accessibility correlated well with RNA expression within cell types, and we demonstrate that more in-depth examination of gene regions can indicate the presence of enhancer elements.

Methods

Workflow of Single-Cell ATAC-Seq and mRNA Whole Transcriptome Analysis on BD Rhapsody™ HT Xpress System

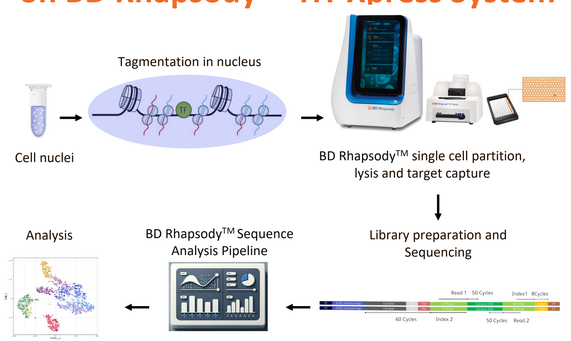

Figure 1. Following the isolation of nuclei from cells, they undergo a tagmentation process utilizing a custom Tn5 transposase. This critical step involves the selective cutting of accessible chromatin regions, during which adapter sequences (pre-loaded onto the Tn5) are concurrently attached to both ends of the resulting fragments. The tagmented nuclei are then precisely loaded into a BD Rhapsody™ 8-Lane Cartridge, where they are individually partitioned into wells equipped with BD Rhapsody™ Enhanced Cell Capture Beads. This process ensures the simultaneous capture of both ATAC and mRNA fragments on the beads. Subsequently, distinct Whole Transcriptome Amplification (WTA) and ATAC libraries are prepared from these beads. These libraries are then sequenced and processed through a comprehensive analysis pipeline, leading to the generation of high-quality data that is fully compatible with existing bioinformatics tools, ready for in-depth analysis.

Summary

Empowering Molecular Biology and Immunology research with an accessible Multiomic scATAC-Seq Assay and analysis pipeline.

- Our study successfully integrates scRNA-Seq and scATAC-Seq through the BD Rhapsody™ Single-cell Analysis System for comprehensive multiomic analysis of PBMCs from multiple donors.
- This gentle approach allows for the simultaneous examination of gene expression and chromatin accessibility, providing a deeper understanding of cellular heterogeneity, development, and gene regulatory mechanisms.
- Our analysis reveals enriched transcription factor motifs and demonstrate a correlation between mRNA profiles and chromatin accessibility regions (CARs).
- Our data demonstrate consistency across both standalone scATAC-Seq and multiomic scATAC-Seq + scRNA-Seq analyses.
- Results of our pipeline are formatted for immediate compatibility with commonly used bioinformatic tools, allowing a quick and easy start to downstream analysis.

Results

Sensitivity and specificity metrics for the BD Rhapsody™ Single-Cell ATAC-Seq Assay

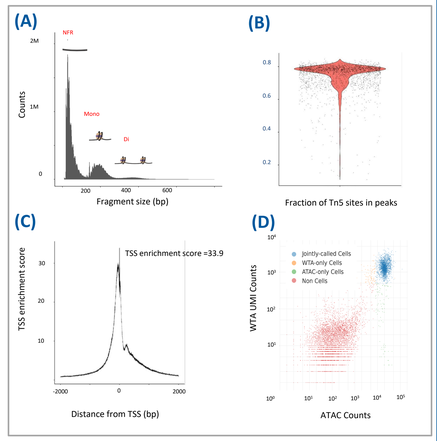

Figure 2. Isolated nuclei form human peripheral blood mononuclear cells (PBMCs) were subjected to scATAC-Seq and mRNA Whole Transcriptome Analysis (WTA) Library Preparation using the BD Rhapsody™ System. **(A)** Fragment size distribution plot, demonstrating a preferential transposition of the nucleosome-free regions of open chromatin (large peak at <147 bp) and further peaks reflecting the ~147 bp repeating pattern of nucleosome positioning. **(B)** Fraction of Tn5 transposase sites from each cell within identified chromatin peaks, highlighting precise chromatin targeting. **(C)** Transcription Start Site (TSS) enrichment plot, displaying the read density around TSSs from the genomic background, emphasizing the accessible state of promoter regions for transcriptional activities. **(D)** Scatter plot of BD Rhapsody™ beads showing cell-calling from a multiomic experiment with human PBMCs. RNA molecules are plotted as a function of transposition events for each bead, and the beads that were called as true cells by RNA, ATAC, and both assays are indicated by color.

RNA and ATAC data can be combined for dimensionality reduction

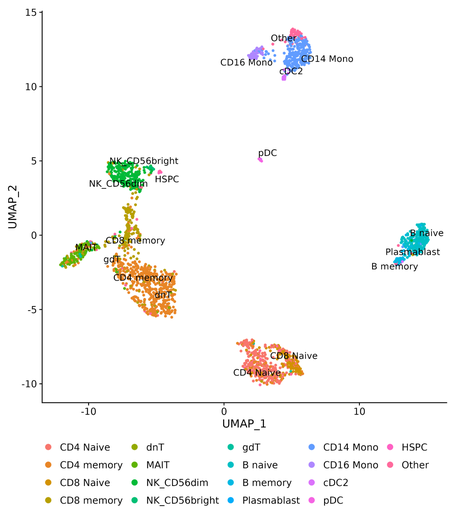

Figure 3. Dimensionality reduction is performed by Uniform Manifold Approximation and Projection (UMAP), utilizing scRNA and scATAC-Seq data jointly. These cells were then annotated as members of distinct cell types using a PBMC reference atlas (Hao 2021) and the R packages Seurat v4.4 and Signac v1.12. These annotations are based on the RNA data, but as high-quality reference atlases become more common, the same approach will be applicable to scATAC-Seq data. Lower-resolution annotations are already possible by inferring gene activity from gene body accessibility (See Figure 5).

Epigenome is dynamic beyond gene boundaries

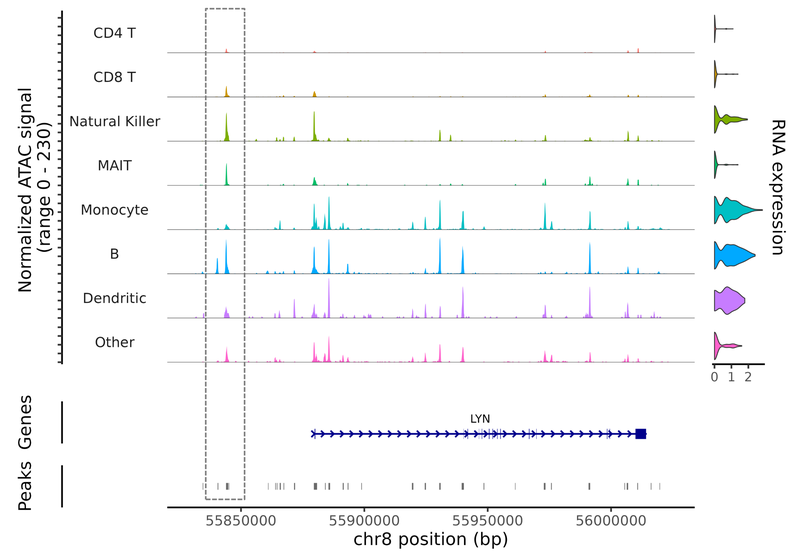

Figure 4. Aggregate ATAC signal within each cell type in the region surrounding the differentially expressed gene LYN. Log-normalized expression of LYN in the cells of each group is displayed in violin plots. Detected peaks of transposase activity are shown below the gene annotation. Note that some obvious peaks are visible upstream of the gene TSS, indicating a likely enhancer element (dotted box).

Transcription factor motif enrichment across cell types

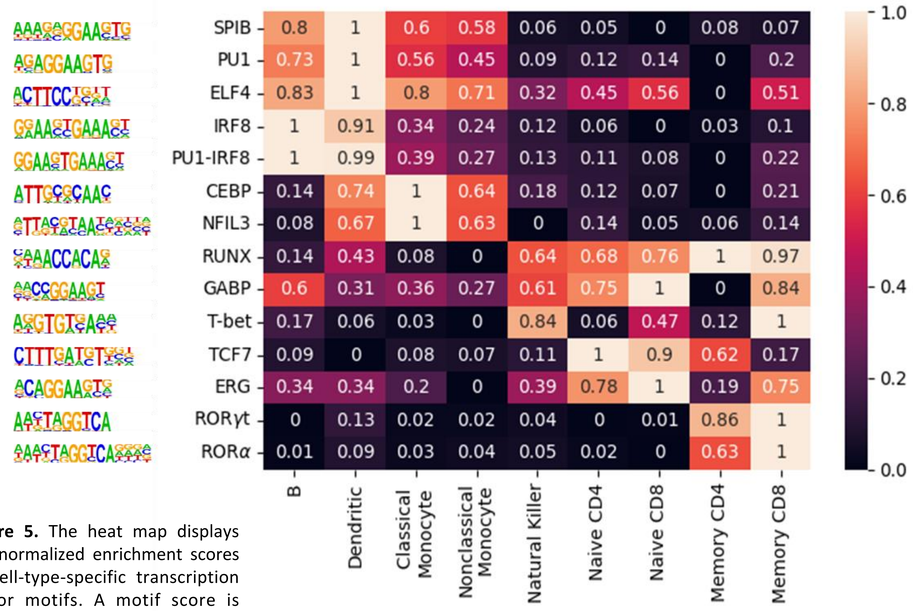

Figure 5. The heat map displays the normalized enrichment scores of cell-type-specific transcription factor motifs. A motif score is calculated based on binomial distribution, using Homer. For each pair of a cell type and a transcription factor, it calculates a relative motif enrichment in the differentially accessible regions of the given cell type, relative to randomly-selected background regions with controlling GC-content. Then the motif scores were normalized across cell types for each motif, with a scale ranging from 0 to 1. A score of 0 represents the least enrichment, while a score of 1 represents the highest enrichment of the motif. Models of transcription factor binding motifs are displayed to the left.

Expression correlates with gene accessibility within cell types

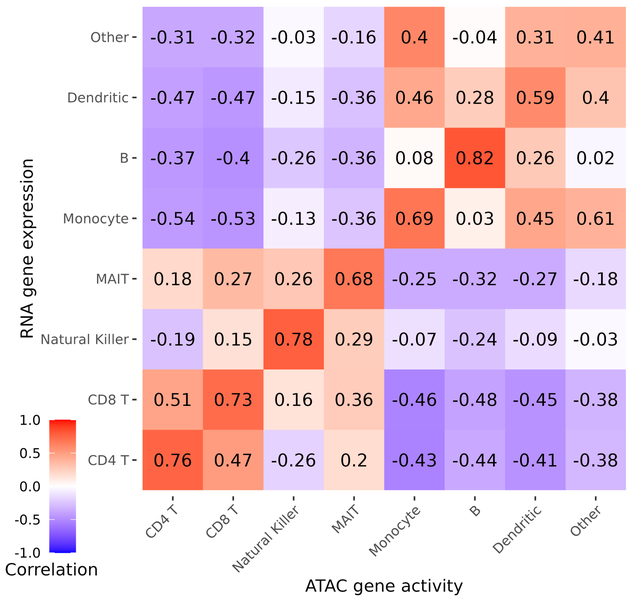

Figure 6. This heatmap represents the Pearson's correlation coefficients between scATAC-Seq gene activity scores and the corresponding gene expression values aggregated by cell type within a human PBMC data set. The Y-axis corresponds to direct RNA-measured gene expression, while the X-axis is inferred "gene activity", based on transposase activity within and 2000bp upstream of annotated gene bodies. There is high correlation between gene accessibility and expression within most of the cell-types. This is not a symmetrical heatmap because comparing the ATAC accessibility of one cell group to another group's RNA expression is not the same as comparing the RNA from the first group with the chromatin accessibility of the second. It is good that correlation is high but not perfect: if correlation were perfect, there would be no point in doing both assays.

High reproducibility across PBMC replicates with BD Rhapsody™ multiomic scATAC-Seq assays: A peak comparison analysis

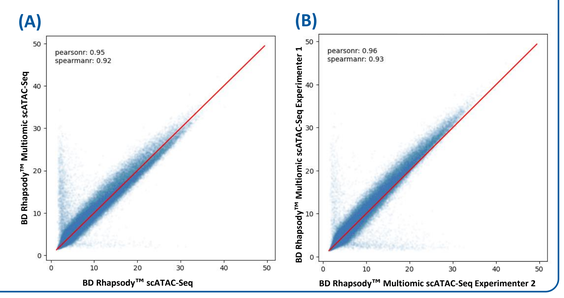

Figure 7. **(A)** Correlation analysis of PBMC sample replicates using standalone and multiomic scATAC-Seq assays demonstrates strong agreement between techniques through peak comparison. The analysis reveals a high correlation, indicating that the WTA assay does not interfere with the ATAC results.
(B) There is also high correlation in the ATAC peaks found between two experiments performed by different scientists on the same PBMC sample, underscoring the reproducibility and reliability of the technique across multiple experiments.